

ON THE SPEED OF CONVERGENCE OF NEWTON'S METHOD FOR COMPLEX POLYNOMIALS

TODOR BILAREV, MAGNUS ASPENBERG, AND DIERK SCHLEICHER

ABSTRACT. We investigate Newton's method for complex polynomials of arbitrary degree d , normalized so that all their roots are in the unit disk. For each degree d , we give an explicit set S_d of $3.33d \log^2 d(1+o(1))$ points with the following universal property: for every normalized polynomial of degree d there are d starting points in S_d whose Newton iterations find all the roots. If the roots are uniformly and independently distributed, we show that the number of iterations for these d starting points to reach all roots with precision ε is $O(d^2 \log^4 d + d \log |\log \varepsilon|)$ (with probability p_d tending to 1 as $d \rightarrow \infty$). This is an improvement of an earlier result in [Sch], where the number of iterations is shown to be $O(d^4 \log^2 d + d^3 \log^2 d |\log \varepsilon|)$ in the worst case (allowing multiple roots) and $O(d^3 \log^2 d (\log d + \log \delta) + d \log |\log \varepsilon|)$ for well-separated (so-called δ -separated) roots.

Our result is almost optimal for this kind of starting points in the sense that the number of iterations can never be smaller than $O(d^2)$ for fixed ε .

1. INTRODUCTION

Newton's root finding method is an old and classical method for finding roots of a differentiable function; it goes back to Newton in the 18th century, perhaps earlier. It was one of the main reasons why A. Douady, J. Hubbard and others in the late 1970s studied iterations of complex analytic functions. The main question was to know where to start the Newton iterative method in order to converge to the roots of the polynomial. Newton's method is known as rapidly converging near the roots (usually with quadratic convergence), but had a reputation that its global dynamics was difficult to understand, so that in practice other methods for root finding were used. See [R] for an overview on recent results about Newton's method.

Meanwhile, some small sets of good starting points are known: there are explicit deterministic sets with $O(d \log^2 d)$ points that are guaranteed to find all roots of appropriately normalized polynomials of degree d [HSS], and probabilistic sets with as few as $O(d(\log \log d)^2)$ points [BLS].

We are interested in the question how many iterations are required until all roots are found with prescribed precision ε . In [Sch], it is shown that among a set of starting points as specified above, there are d points that converge to the d roots and require at most $O(d^4 \log^2 d + d^3 \log^2 d |\log \varepsilon|)$ to get ε -close to the d roots in the worst case; for randomly placed roots (or for roots at mutual distance at least δ for some $\delta > 0$), the required number of iterations is no more than $O(d^3 \log^3 d + d \log |\log \varepsilon|)$ (with the constant depending on δ). This is about one power of d away from the best possible bounds.

In this paper, we show that Newton's method is about as fast as theoretically possible. We consider the space of polynomials of degree d , normalized so as to have all roots in the complex unit disk \mathbb{D} . Our main result is the following.

Theorem 1 (Quadratic Convergence in Expected Case). *For every degree d , there is an explicit universal set \mathcal{S}_d of points in \mathbb{C} , with $|\mathcal{S}_d| = 3.33d \log^2 d(1 + o(1))$, with the following property: suppose that $\alpha_1, \dots, \alpha_n$ are uniformly and independently distributed in the unit disk and put $p(z) = \prod_{j=1}^d (z - \alpha_j)$. Then there are d starting points in \mathcal{S}_d such that with probability $p_d \rightarrow 1$ as $d \rightarrow \infty$, the number of iterations needed to approximate all d roots with precision ε starting at these d points is*

$$O(d^2 \log^4 d + d \log |\log \varepsilon|).$$

Remark 1. As stated, the theorem deals with d distinguishable (i.e., ordered) roots and their associated probability distribution. We prove that the same result holds if we identify our polynomials in terms of their sets of *indistinguishable* roots, as two polynomials $p(z) = \prod_{j=1}^d (z - \alpha_j)$ and $q(z) = \prod_{j=1}^d (z - \beta_j)$ are the same if their unordered sets of roots $\{\alpha_1, \dots, \alpha_d\}$ and $\{\beta_1, \dots, \beta_d\}$ are equal.

Remark 2. This bound on the number of iterations is optimal in the sense that there is no bound on the number of iterations in the same generality that for fixed ε has asymptotics in $o(d^2)$, so we are away from the best possible bound only by a factor of about $O(\log^4 d)$.

2. GOOD STARTING POINTS FOR NEWTON'S METHOD

Studying the geometry of the immediate basins outside the unit disk \mathbb{D} , in [HSS] we proved the existence of a universal starting set with $1.11d \log^2 d$ points depending only on d such that for every polynomial of degree d with all roots in the unit disk, and for every root, there is a point in the set which is in the immediate basin of this root. Enlarging the set by a factor of 3 approximately, in [Sch] we obtained a set of starting points \mathcal{S}_d which ensured that for each polynomial p and each root α there is a point z in \mathcal{S}_d intersecting the immediate basin U of α in the “middle third” of the “thickest” *channel*, where a channel is an unbounded connected component of $U \setminus \mathbb{D}$. Being in this middle third implies an upper bound on the displacement $d_U(z, N_p(z))$ in terms of the Poincaré metric of the immediate basin. It also turns out that the orbit of z under iteration of Newton map does not leave $D_R(0)$, the disk of radius R centered at the origin for some bounded value of R . We will refer to such points as having *R -central orbits*.

More precisely, let \mathcal{S}_d be defined as follows.

Definition 2 (Efficient Grid of Starting Points). For each degree d , construct a circular grid \mathcal{S}_d as follows. For $k = 1, \dots, s = \lceil 0.4 \log d \rceil$, set

$$r_k = (1 + \sqrt{2}) \left(\frac{d-1}{d} \right)^{\frac{2k-1}{4s}},$$

and for each circle around 0 of radius r_k , choose $\lceil 8.33d \log d \rceil$ equidistant points (independently for all the circles).

The set \mathcal{S}_d thus constructed has $3.33(1 + o(1))d \log^2 d$ points. The following theorem is proven in [Sch, Theorem 8].

Theorem 3. *For each degree d , the set \mathcal{S}_d has the following universal property. If p is any complex polynomial, normalized so that all its roots are in \mathbb{D} , then there are d points $z^{(1)}, \dots, z^{(d)}$ in \mathcal{S}_d whose Newton iterations converge to the d roots of p . If α is a root of p and U is the immediate basin of α , then there is an index i such that*

$z^{(i)} \in U$ with $d_U(z^{(i)}, N_p(z^{(i)})) < 2 \log d$. In addition, $z^{(1)}, \dots, z^{(d)}$ have R -central orbits for

$$R \leq 5 \left(\frac{d}{d-1} \right)^{\lceil 5\pi(\log d + 1) \rceil}.$$

For $d = 100$, we have $R < 14$; for $d = 1000$, we have $R < 7.5$; and asymptotically the upper bound on R tends to 7.

The result provides an upper bound for R that is uniform in d . This set of starting points will be the basis for the discussion which follows.

3. UNIFORMLY DISTRIBUTED ROOTS

In this manuscript we investigate the Newton map for complex polynomials with randomly distributed roots. In this section, we fix notation and give the strategy of the proof of our main result, Theorem 1.

Let α be a simple root of the polynomial p of degree d and U be the immediate basin of attraction of α . By the discussion in the previous section, there exists $z_1 \in \mathcal{S}_d$ with R -central orbit in U , i.e. under iteration of the Newton map N_p the orbit converges to α and stays within $D_R(0)$. Let $z_{n+1} := N_p(z_n)$ for $n \geq 1$. For any two consecutive points z_n and z_{n+1} along the orbit of z_1 , in [Sch, Section 4] we constructed “thick” curves that, roughly speaking, “use up” area at least $|z_n - z_{n+1}|^2 / (2\tau)$ with $\tau := d_U(z_1, z_2) < 2 \log d$. In the region of quadratic convergence (near the root α), only $\log_2 |\log_2 \varepsilon - 5|$ iterations are sufficient to get ε -close to the root. Outside this region, two such curves with base points z_n and $z_{n'}$ are disjoint if $n' - n \geq 2\tau + 6$ [Sch, Lemma 11]. The bound $O(d^3 \log^3 d + d \log |\log \varepsilon|)$ follows from lower bounds on the displacements $|z_n - z_{n+1}|$ along the orbit. The main improvement in this paper is on the lower bounds on the displacements when the roots are randomly distributed.

As in [Sch], we partition $D_R(0)$ (the disk of radius R centered at the origin) into domains

$$S_k := \left\{ z \in D_R(0) : \min_j |z - \alpha_j| \in \left(2^{-(k+1)}, 2^{-k} \right] \right\}, \quad k \in \mathbb{Z}.$$

It turns out that if the roots are randomly distributed in the unit disk, then with high probability the following holds: there exists a universal constant C such that for every n we have the following estimates

$$|z_n - z_{n+1}| \geq \begin{cases} \frac{C}{d \log d} & \text{if } z_n \in S_k \text{ with } k \leq \log_2 d; \\ \frac{C}{2^k} & \text{otherwise.} \end{cases}$$

If $z_n \in S_k$ with $k \leq \log_2 d$, then we say that we are “in the far case”, as z_n is far from all the roots. Since each such iteration uses an area of at least $|z_n - z_{n+1}|^2 / (2\tau)$, and one in $2\tau + 6$ such areas are disjoint, the total number of orbit points in the far case is bounded by $O(d^2 \log^4 d)$.

[Sch, Lemma 16] says that if the orbit gets very close to some root in comparison to the other roots, then it has entered the region of quadratic convergence of that root where only $\log_2 |\log_2 \varepsilon - 5|$ are sufficient to approximate it within an ε -neighborhood. We call this “the near case”.

For randomly distributed roots, the mutual distance between roots is large enough so away from the region of quadratic convergence, we only need to consider $k \leq 3 + (2 + \eta) \log_2 d$ for a certain $\eta > 0$. We define the “intermediate case” as those $z_n \in S_k$ with $\log_2 d < k \leq 3 + (2 + \eta) \log_2 d$. Each domain S_k has area $O(d4^{-k})$, and

each iteration with $z_n \in S_k$ uses area about $(C/2^k k)^2/2\tau \approx C^2/4^k k^2 \tau$, so the number of orbit points in the intermediate case is at most $O(dk^2\tau)$ for each k , times the usual factor $2\tau + 6$ to make the areas disjoint. But $\log_2 d < k \leq 3 + (2 + \eta)\log_2 d$ and $\tau = O(\log d)$, so the total number of iterations in the intermediate case is $O(d \log^5 d)$.

In the subsequent sections we will make these arguments precise.

3.1. On the distribution of the roots. In order to get a lower bound on the expected displacement, we will first investigate the distribution of the roots. We will be interested in two different kind of probability spaces. The first space $\mathcal{P}_d = \{(x_1, \dots, x_d) : x_i \in \mathbb{D}\}$ consists of all polynomials with d *distinguishable* roots in the unit disk, normalized so as to have leading coefficients 1, and the probability measure is induced by the Lebesgue measure on \mathbb{D}^d . The second space \mathcal{P}_d/S_d consists of all polynomials with *indistinguishable* roots in the unit disk, i.e. the quotient probability space of the standard action of the symmetric group S_d on \mathcal{P}_d defined by permuting the roots.

The following lemma is the probabilistic ingredient of the main theorem. It certainly isn't new, but easier verified than looked up in the library.

Lemma 4 (Base- d numbers). *Let M_d be the set of all d -digit numbers in base d .*

(a) *The probability that that a randomly chosen number $a \in M_d$ does not have a digit repeating more than $O(\log d)$ times is at least $1 - 1/d$.*

(b) *Let \sim be an equivalence relation on M_d defined as follows: $a \sim b \Leftrightarrow \exists \sigma \in S_d$ with $a = \sigma b$, i.e. two elements are equivalent if they have the same sets of digits counted with multiplicities. Then the probability that a randomly chosen element $[a] \in M_d/\sim$ does not have a digit repeating more than $O(\log d)$ is at least $1 - 1/d$.*

Proof. (a) For fixed i , the number of d -digit numbers which contain at least α digits i is at most $\binom{d}{\alpha} d^{d-\alpha}$. Thus the number of d -digit numbers which contain a symbol repeating at least α times is at most

$$d \binom{d}{\alpha} d^{d-\alpha} < \frac{d}{\alpha!} d^d.$$

So the probability that a randomly selected number in M_d contains at least α identical digits is at most $\frac{d}{\alpha!}$ since $|M_d| = d^d$. Therefore, with probability at least $1 - \frac{d}{\alpha!}$, a randomly selected number in M_d does not have a digit repeating more than α times.

Note that if $\alpha! \geq d^2$ we have $1 - \frac{d}{\alpha!} \geq 1 - \frac{1}{d}$. Therefore, by taking α such that $(\alpha - 1)! < d^2 \leq \alpha!$ (which implies that $\alpha \in O(\log d)$), we prove the first part of the claim.

(b) Note that the elements of M_d/\sim can be bijectively mapped to the set $\text{Mult}_d = \{(x_0, \dots, x_{d-1}) : x_i \in \mathbb{Z}_{\geq 0}, x_0 + \dots + x_{d-1} = d\}$ as follows: for $[a] \in M_d/\sim$ let x_i be the multiplicity of digit i in every $a \in [a]$. It is well known and easy to see that

$$(1) \quad \left| \{(x_0, \dots, x_{r-1}) : x_i \in \mathbb{Z}_{\geq 0}, x_0 + x_1 + \dots + x_{r-1} = n\} \right| = \binom{n+r-1}{r-1}.$$

Thus we have $|\text{Mult}_d| = \binom{2d-1}{d-1}$. On the other hand, the number of elements in Mult_d with first component at least α is equal to the cardinality of

$$\{(x_2, \dots, x_d) : x_i \in \mathbb{Z}_{\geq 0}, x_2 + \dots + x_d \leq d - \alpha\}$$

which has the same cardinality as

$$\{(y_1, x_2, \dots, x_d) : y_1, x_i \in \mathbb{Z}_{\geq 0}, y_1 + x_2 + \dots + x_d = d - \alpha\}.$$

Again by (1) this quantity equals to $\binom{2d-\alpha-1}{d-1}$. Therefore, the number of elements of Mult_d with a component at least α , i.e. the number of elements of M_d/\sim with a digit repeating at least α times, is at most

$$d \binom{2d-\alpha-1}{d-1}.$$

Hence the probability that a number of M_d/\sim has a digit repeating at least α times is at most

$$\begin{aligned} \frac{d \binom{2d-\alpha-1}{d-1}}{\binom{2d-1}{d-1}} &= \frac{d(2d-\alpha-1)!(d-1)!d!}{(2d-1)!(d-1)!(d-\alpha)!} = \\ &= d \frac{(d-\alpha+1)(d-\alpha+2)\cdots d}{(2d-\alpha)(2d-\alpha+1)\cdots(2d-1)} \leq d \left(\frac{1}{2}\right)^{\alpha-1} \frac{d}{2d-1}. \end{aligned}$$

Hence for $\alpha = \lceil 2 \log_2 d + 1 \rceil \in O(\log d)$ the second part of the claim follows. \square

Remark 3. The statement remains true if we replace the probability $1 - 1/d$ by $1 - 1/d^c$ for any constant $c \geq 1$.

If the roots are randomly distributed in the unit disk one should expect that the number of roots in a region is proportional to its area. The previous claim easily implies the following statement.

Lemma 5. *Let a polynomial with roots x_1, \dots, x_d be randomly chosen in \mathcal{P}_d or \mathcal{P}_d/S_d . Then there exists a constant $C_d \in O(\log d)$ such that with probability at least $1 - 1/d$ the following holds true: every disk in \mathbb{C} with area $A = \pi r^2$ contains at most $k(A)$ points among x_1, \dots, x_d with*

$$k(A) = \begin{cases} C_d d A & \text{if } A \geq 1/d; \\ C_d & \text{otherwise.} \end{cases}$$

Proof. Without loss of generality we can assume that $d = (2k+1)^2$ for an integer k (the general case follows by enlarging $C_{(2k+1)^2}$ by a bounded factor). Then the unit disk can be subdivided into d pieces as follows (compare Fig. 1): the first piece is a disk with center 0 and radius $r_0 = 1/\sqrt{d}$; next, consider the annuli A_s bounded between circles around 0 of radii $(2s-1)r_0$ and $(2s+1)r_0$ for $s = 1, \dots, k$, and subdivide each annulus A_s into exactly $8s$ pieces of equal area by drawing $8s$ radial segments. Thus we construct exactly d pieces with equal area and diameters comparable with r_0 . By Lemma 4 it follows that each of the pieces contains at most $O(\log d)$ of the points x_1, \dots, x_d with probability at least $1 - 1/d$ (in both cases of distinguishable and indistinguishable roots): in the case of distinguishable roots, the i -th digit of a d -digit number specifies the number of the piece containing the i -th root; in the other case, the same symmetries apply on both sides of the equality.

Hence, the claim is true for that particular partition of the unit disk. This implies the general claim as follows. It is easy to see that each square of side length at most r_0 in the complex plane can intersect at most a constant number C' of these pieces, where C' does not depend on d . Consider a square S for which the unit circle is inscribed, for example the one with sides parallel to the real and imaginary axes. Subdivide it into d equal squares of side length r_0 (using the fact that d is a square and $r_0 = 1/\sqrt{d}$). Then each of these smaller squares will intersect at most C' pieces from the partition of the unit disk (some squares will not intersect any). Therefore each of the small squares contains at most $C'' \log d$ points for some constant C''

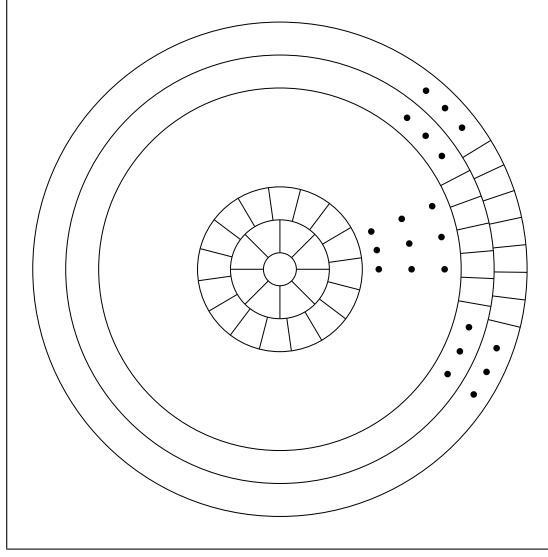


FIGURE 1. Partition of the unit disk into smaller pieces of similar sizes.

which does not depend on d . Since each square of side length r_0 (possibly rotated) intersects at most 9 of these squares dividing S , we conclude that each square of side length r_0 contains, with probability at least $1 - 1/d$, at most $C \log d$ of the points x_1, \dots, x_d for $C = 9C''$. If we group every 4 neighboring small squares (of side length r_0) and repeat the argument, we get that each square of side length $2r_0$ contains at most $4C \log d$ points, and so on for squares of side length $4r_0, 8r_0, \dots$. Thus an arbitrary square of side length $x \in [2^k r_0, 2^{k+1} r_0]$ contains at most $2^{2k+2} C \log d \leq 4C(x^2/r_0^2) \log d \approx 4Cx^2 d \log d$ points since it is contained some square of side length $2^{k+1} r_0$, and thus enlarging the constant by a factor of 4 the lemma will hold true for squares. Since each disk of radius r is contained in a square of side length $2r$, the bound on the number of points in an arbitrary disk follows. \square

Now we prove the following claim about the mutual distance for randomly distributed points in the unit disk.

Lemma 6. *Let the polynomial $p(z)$ be randomly chosen in \mathcal{P}_d or \mathcal{P}_d/S_d . Then the mutual distance between any pair of its roots is at least $1/d^{1+\eta}$, for any fixed $\eta > 0$, with probability at least $1 - 1/d^{2\eta}$.*

Proof. First, note that the claim for a randomly chosen polynomial in \mathcal{P}_d/S_d follows from the claim for a randomly chosen polynomial in \mathcal{P}_d . Choosing randomly a polynomial in \mathcal{P}_d is equivalent to choosing randomly and independently its roots. For a positive number r , the probability $p_{d,r}$ that d uniformly and independently distributed points in the unit disk have mutual distance at least r is at least

$$p_{d,r} \geq (1 - r^2)(1 - 2r^2) \dots (1 - (d-1)r^2)$$

(the unit disk has area π , and after k roots are selected, the $k+1$ -st root must avoid an area of at most $k\pi r^2$; this has probability $(\pi - k\pi r^2)/\pi = 1 - kr^2$).

Since $\log(1+x) \geq x/(1+x)$ for $x > -1$, we get

$$\log p_{d,r} \geq \sum_{k=1}^{d-1} \log(1 - kr^2) \geq \sum_{k=1}^{d-1} \frac{-kr^2}{1 - kr^2} \geq -r^2 \frac{\sum_{k=1}^{d-1} k}{1 - dr^2} \geq -r^2 \frac{d^2/2}{1 - dr^2} \geq -d^2 r^2,$$

where the last inequality holds if $dr^2 < 1/2$. Hence

$$p_{d,r} \geq \exp(-d^2 r^2) \geq 1 - d^2 r^2.$$

If $r = 1/d^{1+\eta}$ (which satisfies $dr^2 < 1/2$), then $p_{d,r} \geq 1 - 1/d^{2\eta}$ and thus the claim follows. \square

Remark 4. The precise value of η in this lemma is not very important to us, as eventually it will only affect constants in the bounds we obtain.

An immediate corollary of the lemma is an upper bound on the distance of z_n to the closest root which guarantees quadratic convergence.

Corollary 7. *If the d roots are randomly chosen and $z_n \in S_k$ with $2^{-k} < 1/8d^{2+\eta}$ and $\eta > 0$, then with probability at least $1 - 1/d^{2\eta}$ the orbit of z_n converges to the closest root α , and $\log_2 |\log_2 \varepsilon - 5|$ iterations of z_n are sufficient to get ε -close to α .*

Proof. Indeed, if $z_n \in S_k$ and α is the closest root to z_n , then $|z_n - \alpha| < 1/8d^{2+\eta}$ and for every root $\alpha_j \neq \alpha$ we have

$$|z_n - \alpha_j| \geq |\alpha - \alpha_j| - |\alpha - z_n| > 1/d^{1+\eta} - 1/8d^{2+\eta} \geq (8d+1)/8d^{2+\eta} > (4d+3)|z_n - \alpha|$$

(under the conditions of Lemma 6). Therefore by [Sch, Lemma 16], we need no more than $\log_2 |\log_2 \varepsilon - 5|$ iterations to get ε -close to α . \square

We combine the previous two lemmas in the following claim.

Lemma 8. *Let a polynomial with roots x_1, \dots, x_d be randomly chosen in \mathcal{P}_d or \mathcal{P}_d/S_d . Then with probability $p_d \geq 1 - O(d^{-2\eta})$ (for fixed $\eta \in (0, 1/2)$), the following two statements simultaneously hold true:*

Area Condition (AC): *There exists $C_d \in O(\log d)$ such that every disk in \mathbb{C} with area $A = \pi r^2$ contains at most $k(A)$ roots among x_1, \dots, x_d with*

$$k(A) = \begin{cases} C_d d A & \text{if } A \geq 1/d; \\ C_d & \text{otherwise.} \end{cases}$$

Distance Condition (DC): *The mutual distance between any pair of roots is at least $1/d^{1+\eta}$.*

Proof. We are interested in $P(AC = \text{true and } DC = \text{true})$, which equals

$$1 - P(AC = \text{false or } DC = \text{false}) \geq 1 - P(AC = \text{false}) - P(DC = \text{false}).$$

By Lemma 5 we have $P(AC = \text{false}) \leq 1/d$, and by Lemma 6 $P(DC = \text{false}) \leq 1/d^{2\eta}$. Hence the claim follows. \square

3.2. Proof of the main theorem. In this section we will use the two conditions **AC** and **DC** to prove Theorem 1. While **DC** guarantees that proximity to a root implies fast convergence (Corollary 7), **AC** gives a lower bound on the displacements along an orbit far away from the roots. More precisely, the following statement holds true.

Lemma 9. *Suppose that the Area Condition in Lemma 8 holds true. If $z_n \in S_K \cap \mathbb{D}_2(0)$, then*

$$|z_n - z_{n+1}| \geq \frac{1}{(1 + 2C_d)2^{K+1} + 16\pi C_d d},$$

where $C_d \in O(\log d)$. If $z_n \notin \mathbb{D}_2(0)$, then $|z_n - z_{n+1}| > 1/d$.

Proof. The fact that $z_n \in S_K$ means that the closest root, say α , is at distance $c/2^K$ for some $c \in (0.5, 1]$, and all the other roots satisfy $|z_n - \alpha_j| \geq c/2^K$.

First suppose that $z_n \in S_K \cap \mathbb{D}_2(0)$. This implies that $K \geq -2$. Let $T_k := \{z \in \mathbb{C} : 2^{-k-1} < |z - z_n| \leq 2^{-k}\}$ for $k = -2, \dots, K$. Then all the roots are contained in $\bigcup_{k=-2}^K T_k$. The Area Condition implies that there exists a constant $C_d \in O(\log d)$ such that the number of roots in T_k is bounded by $\pi C_d d 4^{-k}$ for $\pi 4^{-k} \geq 1/d$, and by C_d otherwise. Thus we have

$$\begin{aligned} \left| \sum \frac{1}{z_n - \alpha_j} \right| &\leq \left| \frac{1}{z_n - \alpha} \right| + \sum_{\substack{\alpha_j \neq \alpha \\ \alpha_j \in T_k}} \left| \frac{1}{z_n - \alpha_j} \right| = \frac{2^K}{c} + \sum_{k=-2}^K \sum_{\substack{\alpha_j \neq \alpha \\ \alpha_j \in T_k}} \left| \frac{1}{z_n - \alpha_j} \right| \\ &\leq \frac{2^K}{c} + \sum_{k=-2}^{\lfloor 0.5 \log_2 \pi d \rfloor} \sum_{\substack{\alpha_j \neq \alpha \\ \alpha_j \in T_k}} \left| \frac{1}{z_n - \alpha_j} \right| + \sum_{k=1+\lfloor 0.5 \log_2 \pi d \rfloor}^K \sum_{\substack{\alpha_j \neq \alpha \\ \alpha_j \in T_k}} \left| \frac{1}{z_n - \alpha_j} \right| \\ &\leq \frac{2^K}{c} + \sum_{k=-2}^{\lfloor 0.5 \log_2 \pi d \rfloor} \frac{2\pi C_d d 4^{-k}}{2^{-k}} + \sum_{k=1+\lfloor 0.5 \log_2 \pi d \rfloor}^K C_d 2^{k+1} \\ &\leq 2^{K+1} + 16\pi C_d d + C_d 2^{K+2}. \end{aligned}$$

Therefore

$$|z_n - z_{n+1}| = \frac{1}{\left| \sum \frac{1}{z_n - \alpha_j} \right|} \geq \frac{1}{(1 + 2C_d)2^{K+1} + 16\pi C_d d}.$$

For the case $z_n \notin \mathbb{D}_2(0)$ we have

$$|z_n - z_{n+1}|^{-1} = \left| \sum \frac{1}{z_n - \alpha_j} \right| < \sum_{\alpha_j} 1 = d,$$

and so $|z_n - z_{n+1}| > 1/d$. □

Corollary 10. *Suppose that the Area Condition in Lemma 8 holds true. Then there is a universal constant C such that the following statements hold for any $z_n \in S_k$.*

- (1) *If $2^{-k} \geq 1/d$, then $|z_n - z_{n+1}| \geq \frac{C}{d \log d}$.*
- (2) *If $1/8d^{2+\eta} \leq 2^{-k} < 1/d$, then $|z_n - z_{n+1}| \geq \frac{C}{k 2^k}$.*

Proof. For $z_n \in \mathbb{D}_2(0)$, Lemma 9 gives

$$|z_n - z_{n+1}| \geq \frac{1}{(1 + 2C_d)2^{k+1} + 16\pi C_d d}.$$

If $2^{-k} \geq 1/d$, i.e., $2^{k+1} \leq 2d$, the denominator is at most $O(C_d d)$, so the displacement is at least $C'/(d \log d)$ for some universal constant C' (since $C_d \in O(\log d)$). On the other hand, if $1/8d^{2+\eta} \leq 2^{-k} < 1/d$ and thus $d < 2^k$, the denominator is at most $O(C_d 2^k)$. In this case, $C_d \in O(\log d) = O(k)$, so the displacement is at least $C''/(k 2^k)$. Therefore the claim follows if we take $C = \min\{C', C''\}$.

Finally, $z_n \notin \mathbb{D}_2(0)$ implies $k < -1$. Again Lemma 9 gives $|z_n - z_{n+1}| > 1/d$, and thus we finish the proof by possibly decreasing the constant C . \square

The final step towards proving our main result is in the following theorem.

Theorem 11. *Let the polynomial $p(z)$ be randomly chosen in \mathcal{P}_d or \mathcal{P}_d/S_d and let (z_n) be an R -central orbit converging to a root α with $d_U(z_0, z_1) \leq \tau$ for $\tau < 2 \log d$. Then with probability $p_d \geq 1 - O(d^{-2\eta})$ (for fixed $\eta \in (0, 1/2)$), the required number of iterations for z_0 to get ε -close to α is*

$$O(d^2 \log^4 d R^2 + \log |\log \varepsilon - 5|).$$

Before proceeding to the proof of this statement, we will outline the main idea. As in [Sch], we construct “thick” curves connecting orbit points z_n and z_{n+1} that use up certain area contained in a bounded domain. Far from the root, two curves corresponding to z_n and z_m are disjoint provided that $|n - m| > 2\tau + 6$. A lower bound on the area of the “thick” curves gives an upper bound on the number of iterations. Also, near the root the orbit enters the domain of quadratic convergence where only a few iterations are sufficient to approximate the root.

More precisely, let $\varphi: U \rightarrow \mathbb{D}$ be the Riemann map with $\varphi(\alpha) = 0$ considered in [Sch, Section 5]. If $|\varphi(z_n)| < 1/2$ (“region of fast convergence”), then according to [Sch, Lemma 11] we need only $\log_2 |\log_2 \varepsilon - 5|$ iterations to get ε -close to the root α . For orbit points with φ -images having absolute values greater than $e^{1/2} - 1$, we can prove the following.

Lemma 12. *For every n with $|\varphi(z_n)| > e^{1/2} - 1$, there are open connected subsets $V_n \subset D_{2R+2}(0)$ with $z_n, z_{n+1} \in \partial V_n$ and $|V_n| \geq |z_n - z_{n+1}|^2 / 2\tau$, having the following property: whenever n and m are such that $\min\{|\varphi(z_n)|, |\varphi(z_m)|\} > e^{1/2} - 1$ and $|n - m| \geq \lceil 2\tau + 6 \rceil$, we have $V_n \cap V_m = \emptyset$.*

Proof. Let $\gamma: [0, s] \rightarrow U$ be the hyperbolic geodesic within U connecting z_n to z_{n+1} . For each $z = \gamma(t)$, let $\eta(t)$ be the Euclidean distance from $\gamma(t)$ to ∂U , and let X_t be the straight line segment (without endpoints) perpendicular to $\gamma(t)$ of Euclidean length $\eta(t)$, centered at $\gamma(t)$. Let $V_n := \bigcup_{t \in (0, s)} X_t$. Then all V_n are open and connected and $z_n, z_{n+1} \in \partial V_n$, and the area of V_n is at least $|z_n - z_{n+1}|^2 / 2\tau$: this follows as in [Sch, Lemma 9] (in this reference, the areas restricted to certain domains S_k are calculated; omitting this restriction, we obtain the result we need, and the computations only get simpler). Moreover, the orbit (z_n) is R -central and the unit disk contains other roots than α , and hence the length of $\gamma(t)$ for $t \in [0, s]$ is bounded by $R + 1$. This implies that all pieces V_n are contained in $D_{2R+2}(0)$ by construction.

The fact that $V_n \cap V_m$ are disjoint when $|n - m| > 2\tau + 6$ is proved in [Sch, Lemma 12] (again for restricted domains, but this is immaterial for the proof). \square

Proof of Theorem 11. We only need to consider iteration points whose images under φ have absolute values at least $e^{1/2} - 1$. Also, by Lemma 8, the conditions **AC** and **DC** hold true with probability $p_d \geq 1 - O(d^{-2\eta})$.

Choose M so that $2^M - 1 > 2R + 2$. We distinguish the following three cases.

The Far Case: we have $z_n \in S_k$ with $2^{-k} \geq 1/d$. By Corollary 10 (1) we have $|z_n - z_{n+1}| \geq \frac{C}{d \log d}$. Lemma 12 says that any Newton iteration $z_n \mapsto z_{n+1}$ with $z_n \in S_k$ needs area at least

$$\frac{|z_n - z_{n+1}|^2}{2\tau} \geq \frac{C^2}{2\tau d^2 \log^2 d}.$$

Moreover, the pieces of area for the iterations $z_n \mapsto z_{n+1}$ and $z_{n'} \mapsto z_{n'+1}$ are disjoint provided that $n - n' \geq 2\tau + 6$, and all these pieces of area are contained in the disk $D_{2R+2}(0)$ with R universally bounded.

The total number of such iterations $D_{2R+2}(0)$ can accommodate is thus at most

$$C' d^2 (\log d)^2 \tau [2\tau + 6] R^2$$

for a universal constant C' .

The Intermediate Case: we have $z_n \in S_k$ with $1/8d^{2+\eta} \leq 2^{-k} < 1/d$. Then $\log_2 d < k \leq 3 + (2 + \eta) \log_2 d$. By Corollary 10 (2) we have $|z_n - z_{n+1}| \geq C/k2^k$. Thus by [Sch, Proposition 13], the set S_k contains at most

$$\begin{aligned} & \pi d \left(2^{-k+1} + \frac{C}{k2^k} \right)^2 \left(2\tau + 2^{k-1} \frac{C}{k2^k} \right) [2\tau + 6] \frac{k^2 2^{2k}}{C^2} \\ &= \pi d 2^{-2k} k^{-2} (2k + C)^2 \left(2\tau + \frac{C}{2k} \right) [2\tau + 6] \frac{k^2 2^{2k}}{C^2} \\ &= \pi d C^{-2} (2k + C)^2 \left(2\tau + \frac{C}{2k} \right) [2\tau + 6] \\ &\leq \pi d C^{-2} (6 + (4 + 2\eta) \log_2 d + C)^2 \left(2\tau + \frac{C}{2 \log_2 d} \right) [2\tau + 6] \\ &\leq C'' d \log^2 d (2\tau + 1) [2\tau + 6] \end{aligned}$$

orbit points for some universal constant C'' . There are $3 + (1 + \eta) \log d$ possible values of k in the Intermediate Case, so $\bigcup_k S_k$ (for all k in the Intermediate Case) can accommodate at most

$$(1 + \eta) C'' d \log^3 d (2\tau + 1) [2\tau + 6]$$

orbit points for some universal constant C'' .

The Near Case: we have $z_n \in S_k$ with $2^{-k} < 1/8d^{2+\eta}$. By Corollary 7, α is the closest root to z_n and we need $\log_2 |\log_2 \varepsilon - 5|$ iterations to get ε -close to it.

Since $\tau \in O(\log d)$ and the Far Case dominates the Intermediate Case, the claim follows. \square

We now conclude the main statement.

Proof of Theorem 1. By Theorem 3, for each root there is a starting point satisfying the conditions of the theorem. In particular, these orbits are R -central for a universally bounded value of R . Note that the d roots have to compete for the available area in $D_{2R+2}(0)$. Since the estimates in the proof of Theorem 11 are based on the area (except for the Near Case where the orbit gets to the region of quadratic convergence), we get the same estimate for the combined number of iterations (except that the estimate $\log |\log \varepsilon|$ applies for each root separately, thus it is multiplied by d). \square

Remark 5. This result is close to optimal in the sense that the power of d cannot be reduced for our universal set of starting points that is bounded away from the unit disk. The reason is that outside the unit disk N_p is conjugate to the linear map $w \mapsto \frac{d-1}{d}w$ by [HSS, Lemma 4], so at least $O(d)$ iterations are required for each

“good“ starting point to get close to the unit disk where the roots are located, and at least $O(d^2)$ for all the d starting points combined.

REFERENCES

- [BLS] Béla Bollobás, Malte Lackmann, and Dierk Schleicher, *A small probabilistic universal set of starting points for finding roots of complex polynomials by Newton's method*. Mathematics of computation, to appear. arXiv:1009.1843.
- [HSS] John Hubbard, Dierk Schleicher, and Scott Sutherland, *How to find all roots of complex polynomials by Newton's method*. Inventiones Mathematicae **146** (2001), 1–33.
- [R] Johannes Rückert, *Rational and Transcendental Newton Maps*. In: *Holomorphic dynamics and renormalization* (ed. Mikhail Lyubich and Mikhail Yampolski), Fields Inst. Commun. **53**, Amer. Math. Soc., Providence, RI, 2008, pp. 197–211.
- [Sch] Dierk Schleicher, *On the efficient global dynamics of Newton's method for complex polynomials*. Manuscript, submitted (2011). arXiv:1108.5773.